

本周工作：

1. DNN-based data fusion:

- a) Neural network based on back-propagation. But does not work well when the neural network has many hidden layers.
- b) autoencoder and Restricted Boltzmann Machines 可用于分类和信息检索
- c) eg : video and audio 的例子 :
 - i. cross modality learning 只输入一个从而对整体进行调整
 - ii. shared representation learning 联合学习，两个模型进行互补
 - iii. Mult-modal fusion 多模式数据融合
- d) Boltzmann machines : 三个标准：保持概念一致，关联容易找到用于填充 miss 的 data，有利于检索。
- e) DNN 方式效果好坏主要取决于参数的取值

2. Semantic meaning based data fusion

- a) Understand the insight of each dataset and relation between features across different datasets.
- b) mult-view based data fusion
 - i. 不同数据集或不同特征集看成对一个物体不同的角度 view，不同 dataset 中的既有潜在的一致信息，也可以互补。
 - ii. cotrain : 对两者数据共识最大化
 - iii. multiple kernel : combine linear or nor
 - iv. subspace : Obtain a latent subspace shared by multiple views.

3. VAUD :

- a) 时间地点选择后数据存储
- b) 放大缩小后节点大小并调整新建节点时出现的错位 bug
- c) query 时对于条件的判断从而选择进行哪一种查询

附录：

DNN (深层神经网络) Deep Neural Networks : 深层次的神经网络。这个深度网络从结构上讲与传统的多层感知机没有什么不同，并且在有监督学习时算法也是一样的。唯一的区别是这个网络在做有监督学习前要先做非监督学习 $P(x)$ ，然后将非监督学习学到的权值当作有监督学习的初值进行训练 $P(Y|X)$ 。这种学习思路相对于单纯的有监督学习而言有助于降低过拟合的风险，因为它不仅学习了条件概率分布 $P(Y|X)$ ，还学习了 X 和 Y 的联合概率分布。